Secure Localization for Wireless Sensor Networks using Range-Independent Methods

Loukas Lazos and Radha Poovendran

Network Security Lab, Electrical Engineering Department, University of Washington, Seattle, WA, 98195
llazos@u.washington.edu, rp3@u.washington.edu

1 Introduction

Wireless Sensor Networks (WSNs) are envisioned to be integrated into our everyday lives, enabling a wealth of commercial applications such as environmental and habitat monitoring, disaster relief and emergency rescue operations, patient monitoring, as well as military applications such as target detection and tracking. These applications are facilitated by the collaborative processing of the physical properties monitored by the sensors, such as temperature, light, sound, humidity, vibration, acceleration, or air quality.

For most applications of WSNs, knowledge of the origin of the sensed information is critical for taking appropriate action based on the observations. As an example, if a smoke detector reports the break out of a fire, this information, while useful, is not sufficient to initiate proper action. On the other hand, associating the report from the smoke detector in space, enables the timely response to the reported event. Hence, the association of the observations reported by sensors in space increases the quality of the information aggregated via the sensor network. Furthermore, location is assumed to be known in many network operations such as routing protocols where a family of geographically-aided algorithms have been proposed [2], or security protocols where location information is used to prevent threats against network services [13, 16]. In WSNs, enabling sensors to associate their reports with space is achieved via the location estimation process also known as *localization*.

The majority of the localization techniques that are proposed for WSNs, [4, 12, 25, 27, 31, 34] are designed to operate in a benign environments with no security threats. However, WSNs may be deployed in hostile environments and operating unsupervised, and hence, are vulnerable to conventional and novel attacks [11, 30] aimed at interrupting the functionality of location-aware applications by exploiting the vulnerabilities of the localization scheme.

In this chapter, we study the problem of *enabling nodes of a WSN to determine their location even in the presence of malicious adversaries.* This problem will be referred to as *Secure Localization*. We consider secure localization in the context of

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2006		**00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Secure Localization for Wireless Sensor Networks using Range-Independent Methods	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Washington,Department of Electrical Engineering,Seattle,WA,98195**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**		**30**	

the following design goals: (a) decentralized implementation, (b) resource efficiency, and (c) robustness against security threats.

We illustrate a series of attacks against localization schemes for WSNs [11, 13, 26, 28] and propose *SeRLoc*, a robust location estimation scheme for WSNs that achieves decentralized, resource-efficient sensor localization even in the presence of adversaries. We also propose a high resolution localization algorithm called *HiRLoc*, that improves the localization accuracy at the expense of more complicated hardware. Since sensors are hardware and power limited, SeRLoc and HiRLoc rely on a two-tier network architecture. The network consists of a small number of nodes equipped with known coordinates and orientation we call *locators* and a large number of resource-constrained sensor devices with unknown location.

Moreover, since distance measurements are susceptible to distance enlargement/reduction [5], we do not use any such measurements to compute the sensor location. Instead sensors rely on beacon broadcasts from the locator containing localization information to infer their location. We refer to methods that are not using distance measurements as range-independent localization schemes [4, 12, 25]. Methods for securing range-dependent localization schemes are presented in [5, 7].

Since range independent schemes do not rely on any distance measurements to estimate location, they are not vulnerable to range-alteration attacks. However an adversary may launch relay type of attacks such as the wormhole attack [13, 28], impersonation attacks such as the Sybil attack [11, 26], or compromise network entities. First, we describe the impact of these attacks on the location estimation process, and then, we provide mechanisms that allow each sensor to determine its location *even* in the presence of these threats. Furthermore, we analytically evaluate the probability of success for each type of attack using *spatial statistics* theory [9].

The remainder of the chapter is organized as follows. In Section 2 we illustrate different attacks against range-independent location estimation schemes. In Section 3, we state our network model. In Section 4 we describe two algorithms for robustly estimating the position of sensors. In Section 5, we present a threat analysis. In Section 6, we evaluate the performance of SeRLoc and HiRLoc. In Section 7, we present related work and open problems. Section 8 presents our conclusions.

2 Attacks on Range-independent Localization Schemes

In this section we first define the adversarial model considered for WSNs. We then illustrate different types of attacks against range-independent localization schemes.

2.1 Adversarial Model

We assume that the adversary's goal is to mislead sensors to falsely estimate their location. We also assume that in its effort to mislead the sensors, the adversary must remain undetected. We do not consider Denial-of-Service (DoS) attacks against the localization scheme. Such attacks can be easily detected, since sensors will not be

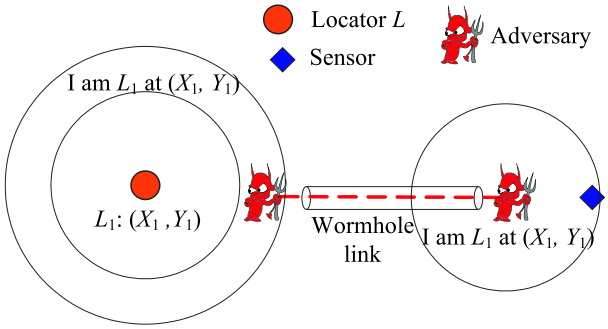


Fig. 1. The adversary records the broadcast of reference point L_1, tunnels it to the region of the sensor under attack and replays it. The sensors believes it is within range of L_1.

able to compute their position. We also do not address attacks against the physical medium such as frequency jamming. Spread spectrum [38] and coding [39] are known to be efficient mechanisms to shield the physical layer against jamming attacks. Also, we do not consider any attack against the Medium Access Control (MAC) protocol that may lead to a denial-of-service (DoS). Secure location estimation schemes that take into account jamming are presented in [5, 20].

2.2 Attack Models

In range-independent location estimation methods, nodes rely on localization information included in beacons transmitted from reference points in order to estimate their position. In order to bias the location estimation process, the adversary attempts to inject bogus localization information into the network. This can be achieved by performing a wormhole (relay) attack [13, 28, 30], an impersonation (Sybil) attack [11, 26], or compromise of reference points. In any of those attacks we assume that at least some valid information not altered by the adversary is present, that allows the node to estimate its position. We now discuss the different attacks against range-independent localization schemes in more detail.

The Wormhole (Relay) Attack

The wormhole attack is a relay type of attack where an adversary relays information transmitted at one part of the network to some distant part of the network, thus violating the geometry of the network and the communication range constraint. To mount a wormhole attack, the adversary initially establishes a direct link referred to as a *wormhole link* between two points in the network. Once the wormhole link is established, the adversary eavesdrops (records) messages at one end of the link, referred to as the *origin point*, tunnels them through the wormhole link and replays them at the other end, referred to as the *destination point*. The wormhole attack is very difficult to detect, since it is launched without compromising any host, or the integrity and authenticity of the communication [13, 28].

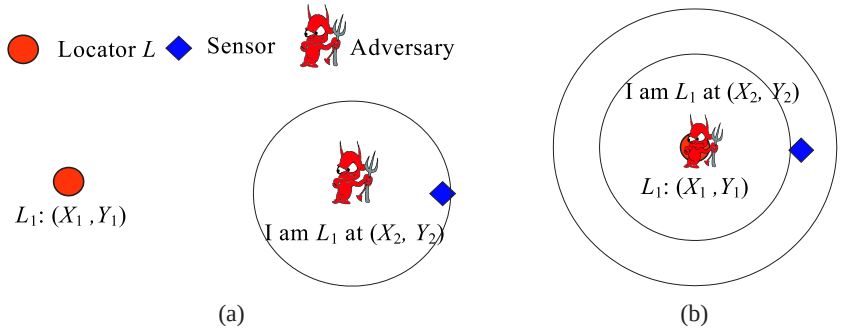


Fig. 2. (a) The adversary impersonates reference point L_1 to a sensor under attack. The sensor is misled to believe it is within range of L_1 with L_1 located at (X_2, Y_2), (b) reference point L_1 is compromised and falsely reports its location.

When an adversary launches a wormhole attack against the location estimation process, sensors located at the destination point of the attack hear beacons transmitted from reference points located at the origin point of the attack. Hence, sensors are misled to believe they are within proximity of reference points at the origin point of the attack. The bogus localization information is properly authenticated by the sensors (since beacons are indeed authentic) and can significantly bias the location estimation at each sensor under attack.

One mechanism for detecting relay type of attacks, is synchronizing the nodes of the network and timestamping each message [13]. Every recipient of a message compares the timestamp with the time when the message is received to determine whether the message has traveled a distance longer than the communication range of the sender. However, when the RF medium is used for transmitting beacons synchronization has to be achieved with nanosecond accuracy [13]. Using a slower medium such as the acoustic medium to transmit beacons to avoid the tight synchronization requirement, leaves the system vulnerable to wormholes when the adversary uses an RF wormhole link to relay the localization information in a timely manner.

2.3 The Impersonation (Sybil) Attack

In the impersonation attack, the adversary assumes one or multiple identities from network nodes and impersonates those nodes to other entities within the network [11, 26]. With respect to the localization process, the adversary impersonates reference points and injects bogus localization information into the network. Unlike the wormhole attack, in the Sybil attack model, the adversary must compromise cryptographic quantities necessary to prove its impersonated IDs to the nodes under attack. Hence, nodes properly authenticate an adversary as a trustable source.

In Figure 2, the adversary impersonates locator L_1 to a sensor that is not within the range of L_1. The sensor under attack is misled to believe that it can hear locator L_1 located at coordinates (X_2, Y_2). The adversary can modify the coordinates contained within the beacon to any arbitrary position within the network.

2.4 Compromise of Network Nodes

The adversary may be also able to compromise network nodes used in the location estimation process and force them to misbehave. For example the adversary may compromise reference points and force them to falsely report their positions. Under node compromise, we assume that the adversary gains full control over the behavior of the entity that has been compromised. This assumption is significantly stronger than the assumption made for launching an impersonation attack where the adversary can only impersonate a node and not alter its behavior (controls only the impersonators).

We assume that the sensors have to receive at least some localization information from uncompromised reference points in order to perform any kind of robust location estimation. In Figure 2(b), we show the compromise of locator L_1 and the broadcast of bogus localization information. The sensor is misled to believe that locator L_1 is located at position (X_2, Y_2).

3 Network Model

In this section, we state our network model assumptions for building our secure location estimation algorithm.

Network Setup: We assume a two-tier network architecture where a set of sensors S of unknown location is randomly deployed with a density ρ_s within an area $\mathcal{A}$, and a set of reference points L we call *locators*, with known location[1] and orientation, also randomly deployed with a density ρ_L.

Antenna Model: We assume that sensors are equipped with omnidirectional antennas and transmit with power P_s, while locators are equipped with M directional antennas with directivity gain $G \rangle 1$, and transmit with power $P_L \rangle P_s$. Since the locator transmission power is higher than the sensor transmission power, the locator-sensor communication channel is asymmetric. For the rest of the chapter, we denote the sensor-to-locator communication range as r, and the locator-to-sensor communication range as R.

System Parameters: Since both locators and sensors are randomly and independently deployed, it is essential to select the system parameters so that sufficient number of locators can communicate with sensors. The random deployment of the locators with a density $\rho_L = \frac{|L|}{\mathcal{A}}$ ($|\cdot|$ denotes the cardinality of a set) is equivalent to a sequence of events following a *homogeneous Poisson point process* of rate ρ_L [9]. The random deployment of sensors with a density $\rho_s = \frac{|S|}{\mathcal{A}}$, is equivalent to a random sampling of the area $\mathcal{A}$ with rate ρ_s [9]. Making use of *Spatial Statistics*

[1] We presume that locators acquire their position either through manual insertion or through GPS receivers [36]. Though GPS signals can be spoofed, knowledge of the coordinates of several nodes is essential to achieve any kind of node localization for any localization scheme.

theory [9], if LH_s denotes the set of locators heard by a sensor s, that is, within range R from s, the probability that s hears exactly k locators, given that the locators are randomly and independently deployed, is given by the Poisson distribution:

$$P(|LH_s| = k) = \frac{(\rho_L \pi R^2)^k}{k!} e^{-\rho_L \pi R^2}. \quad (1)$$

Based on (1) and the independent deployment of sensors, the probability for *every* sensor to hear at least k locators $P(|LH_s| \rangle k)$:

$$P(|LH_s| \geq k, \forall s \in S) = (1 - \sum_{i=0}^{k-1} \frac{(\rho_L \pi R^2)^i}{i!} e^{-\rho_L \pi R^2})^{|S|}. \quad (2)$$

Equation (2) allows the choice of $\rho_L,\ R$ so that a sensor hears at least k locators with any desired probability.

4 Secure Location Estimation in WSN

In this section we describe two location estimation schemes. We first present the SEcure Range-independent LOCalization scheme (*SeRLoc*) that enables sensors to determine their location based on beacon information transmitted by the locators, even in the presence of security threats. We then present the HIgh-resolution LOCalizaion scheme (*HiRLoc*) that improves the location resolution.

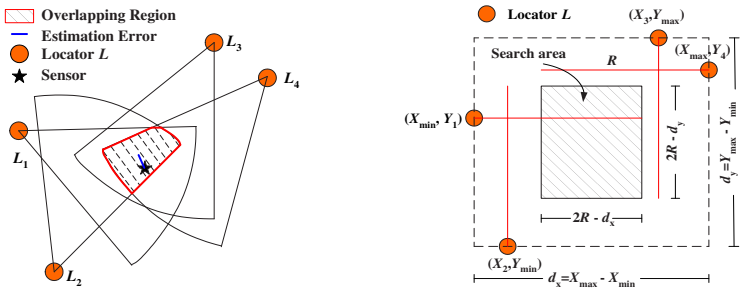
4.1 Location Determination in SeRLoc

In SeRLoc, sensors determine their location based on the localization information included in beacons transmitted by the locators. Figure 3(a) illustrates the idea behind SeRLoc. Each locator transmits beacons at each antenna sector containing (a) the locator's coordinates and, (b) the angles of the antenna boundary lines with respect to a common global axis.

For each locator L_i heard at a sensor s, sensor s defines the sector S_i corresponding to the transmission of that locator where s has to be included. Combining information from multiple locators it defines the *Region Of Intersection* (ROI), as the region where the maximum number of sectors overlap:

$$ROI = \bigcap S_i. \quad (3)$$

The sensor s determines its location as the center of gravity (CoG) of the ROI. The CoG is the least square error solution given that a sensor can lie with equal probability at any point of the ROI. In Figure 3(a), the sensor hears beacons from locators $L_1 \sim L_4$ and determines its position as the CoG of the ROI. We now present the algorithmic details of SeRLoc.



(a) (b)

Fig. 3. (a) The sensor hears locators $L_1 \sim L_4$ and estimates its location as the Center of Gravity (CoG) of the region of intersection. (b) Determination of the search area.

- *Step 1:* **Collection of localization information:** In Step 1, the sensor collects information from all the locators that it can hear. A sensor s can hear all locators $L_i \in L$ that lie within a circle of radius R, centered at s.

$$LH_s = \{L_i : \|s - L_i\| \leq R, \quad L_i \in L\}. \tag{4}$$

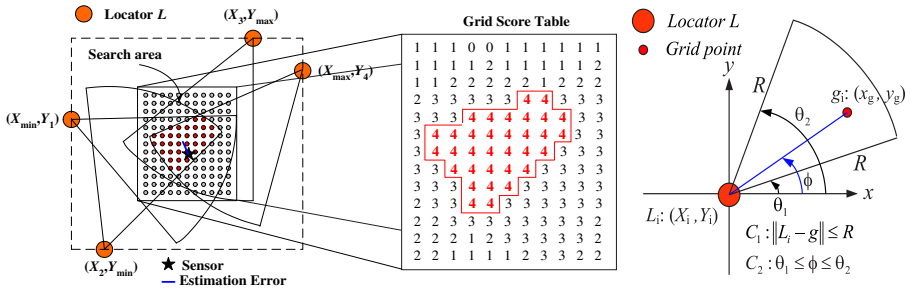
- *Step 2:* **Search area:** In Step 2, the sensor computes a search area for its location. Let $X_{min}, Y_{min}, X_{max}, Y_{max}$ denote the minimum and the maximum locator coordinates form the set LH_s.

$$X_{min} = \min_{L_i \in LH_s} X_i,\ X_{max} = \max_{L_i \in LH_s} X_i,\ Y_{min} = \min_{L_i \in LH_s} Y_i,\ Y_{max} = \max_{L_i \in LH_s} Y_i \tag{5}$$

Since every locator of set LH_s needs to be within a range R from sensor s, if s can hear locator L_i with coordinates (X_{min}, Y_i), it has to be located *left* of the vertical boundary of $(X_{min} + R)$. Similarly, s has to be located *right* of the vertical boundary of $(X_{max} - R)$, *below* the horizontal boundary of $(Y_{min} + R)$, and *above* the horizontal boundary of $(Y_{max} - R)$. The dimensions of the rectangular search area are $(2R - d_x)$x$(2R - d_y)$ where d_x, d_y are the horizontal distance $d_x = X_{max} - X_{min} \leq 2R$ and the vertical distance $d_y = Y_{max} - Y_{min} \leq 2R$, respectively. In Figure 3(b), we show the search area for the network setup in Figure 3(a).

- *Step 3:* **Overlapping region-Majority vote**: In Step 3, sensors determine the ROI of all sectors they hear. Since it would be computationally expensive for each sensor to analytically determine the ROI based on the line intersections, we employ a grid scoring system that defines the ROI based on majority vote.

Grid score table: The sensor places a grid of equally spaced points within the rectangular search area as shown in Figure 4(a). For each grid point, the sensor holds a score in a grid score table, with initial values equal to zero. For each grid point, the sensor executes the *grid-sector test* detailed in the following, to decide if the grid point is included in a sector heard by a locator of set LH_s. If the grid score test is positive the sensor increments the corresponding grid score table value by one, otherwise the value remains unchanged. This process is repeated for all locators



(a) (b)

Fig. 4. (a) Steps 3,4: Placement of a grid of equally spaced points in the search area, and the corresponding grid score table. The sensor estimates its position as the centroid of all grid points with the highest score, (b) Step 3: Grid-sector test for a point g of the search area.

heard LH_s, and all the grid points. The ROI is defined by the grid points that have the highest score in the grid score table. In Figure 4(a), we show the grid score table and the corresponding ROI.

Note that due to the finite grid resolution, error is induced in the calculation. The resolution of the grid can be increased to reduce the error at the expense of energy consumption due to the increased processing time.

Grid-sector test: A point $g : (x_g, y_g)$ is included in a sector of angles $[\theta_1, \theta_2]$ originating from locator L_i if it satisfies two conditions:

$$C_1 : \ \|g - L_i\| \leq R, \qquad C_2 : \ \theta_1 \leq \phi \leq \theta_2, \tag{6}$$

where ϕ is the slope of the line connecting g with L_i. Note that the sensor *does not have to* perform any angle-of-arrival (AOA) measurements. Both the coordinates of the locators and the grid points are known, and, hence the sensor can analytically calculate ϕ. In Figure 4(b), we illustrate the grid-sector test with all angles measured with reference to the x axis.

- *Step 4:* **Location estimation**: The sensor determines its location as the centroid of all the grid points that define the ROI.

$$\tilde{s} : (x_{est}, y_{est}) = \left(\frac{1}{n} \sum_{i=1}^{n} x_{g_i}, \frac{1}{n} \sum_{i=1}^{n} y_{g_i} \right), \tag{7}$$

where n is the number of grid points of the overlapping region, and (x_{g_i}, y_{g_i}) are the coordinates of the grid points.

4.2 HiRLoc: High-resolution Range-Independent Localization Scheme

In this section, we present the High-resolution Range-independent Localization scheme (*HiRLoc*) that allows sensors to determine their location with higher accuracy compared to SeRLoc at the expense of more complex hardware at the locator side.

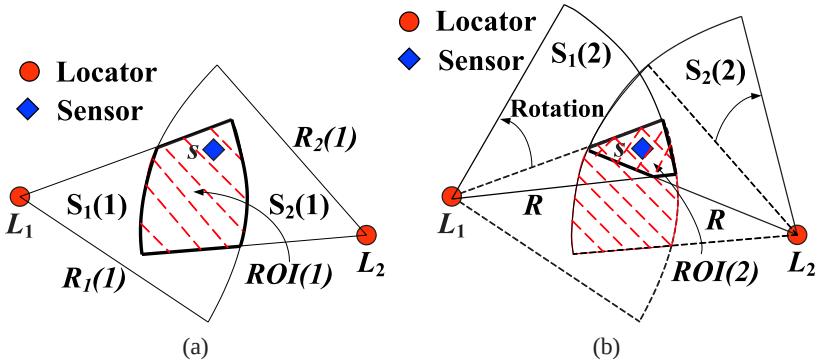


Fig. 5. (a) The sensor is located within the intersection of the sectors $S_1(j), S_2(j)$, which defines the region of intersection ROI. (b) The ROI is reduced by the rotation of the antenna sectors by some angle α.

4.3 Location Determination in HiRLoc

In HiRLoc, localization accuracy in improved by having locators either rotate their antenna system, or change their communication range in order to define new sectors where transmission takes place. Superimposing the sectors indicated by the beacons not only in space but also in time provides the extra location resolution. Based on the beacon information the sensors define the sector area $S_i(j)$ as the confined area covered by the j^{th} transmission of a locator L_i.

By collecting beacons from the locators $L_i \in LH_s$, the sensor can compute its location as the ROI of all the sectors $S_i(j)$. Note that a sensor can hear beacons from multiple locators, and multiple beacons generated by the same locator. Hence, the ROI after the m^{th} round of beacon transmissions can be expressed as the intersection of all the sectors corresponding to the beacons available at each sensor:

$$ROI(m) \overset{\text{(i)}}{=} \bigcap_{i=1}^{|LH_s|} \bigcap_{j=0}^{m} S_i(j) \overset{\text{(ii)}}{=} \bigcap_{j=0}^{m} \left(\bigcap_{i=1}^{|LH_s|} S_i(j) \right), \tag{8}$$

Since the ROI indicates the confined region where the sensor is located, reducing the size of the ROI leads to an increase in the localization accuracy. Based on equation (8), we can reduce the size of the ROI by, (a) reducing the size of the sector areas $S_i(j)$ and, (b) increase the number of intersecting sectors $S_i(j)$.

In HiRLoc, reduction of the ROI is achieved by exploiting the temporal dimension. The locators provide different localization information at consecutive beacon transmissions by, (a) varying the direction of their antennas and, (b) varying the communication range of the transmission via power control. We now explore how both these methods lead to the reduction of the ROI.

1. Varying the antenna orientation: The locators are capable of transmitting at all directions (omnidirectional coverage) using multiple directional antennas. Every antenna has a specific orientation and hence corresponds to a fixed sector area

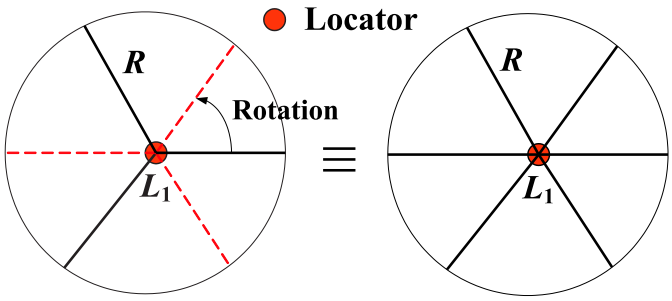


Fig. 6. Locator L_1 is equipped with three directional antennas of beamwidth $\frac{2\pi}{3}$ each. The transmission of beacons at each sector, followed by antenna rotation by $\frac{\pi}{3}$, followed by a transmission of update beacons, is equivalent to equipping L_1 with six directional antennas of beamwidth $\frac{\pi}{3}$.

$S_i(j)$. The antenna orientation is expressed by the angle information contained in the beacon $\theta_i(j) = \{\theta_{i,1}(j), \theta_{i,2}(j)\}$, where $\theta_{i,1}(j), \theta_{i,2}(j)$ denote the lower and upper bounds of the sector $S_i(j)$.

Instead of reducing the size of the intersecting sectors by narrowing the antenna beamwidth, locators can change the orientation of their antennas and re-transmit beacons with the new sector boundaries. A change in the antenna orientation can occur either by changing the orientation of the locators, or by rotation of their antenna system. A sensor collects multiple sector information from each locator over a sequence of transmissions: $S_i(j) = S_i(\theta_i(j), j), j = 1 \ldots Q$. As expressed by equation (8), the intersection of a larger number of *distinct* sectors leads to a reduction in the size of the ROI. As an example, consider Figure 5 where a sensor s hears locators L_1, L_2. In Figure 5(a), we show the first round of beacon transmissions by the locators L_1, L_2, and the corresponding $ROI(1)$. In Figure 5(b), the locators L_1, L_2 rotate their antennas by an angle α and transmit the second round of beacons with the new sector boundaries. The ROI in the two rounds of beacon transmissions, can be expressed as:

$$ROI(1) = S_1(1) \cap S_2(1) \qquad ROI(2) = ROI(1) \cap S_1(2) \cap S_2(2). \tag{9}$$

The antenna rotation over time can be interpreted as an increase on the number of antenna sectors of each locator via superposition over time. For example, consider Figure 6, where a locator is equipped with three directional antennas of beamwidth $\frac{2\pi}{3}$. Transmission of one round of beacons, followed by antenna rotation by $\frac{\pi}{3}$ and re-transmission of the updated beacons is equivalent to transmitting one round of beacons when locators are equipped with six directional antennas of beamwidth $\frac{\pi}{3}$.

2. Varying the Communication range: A second approach to reduce the area of the ROI, is to reduce the size of the intersecting sectors. This can be achieved by allowing locators to decrease their transmission power and re-broadcast beacons with the new communication range information. In such a case, the sector area $S_i(j)$ is dependent upon the communication range $R_i(j)$ at the j^{th} transmission, i.e. $S_i(j) =$

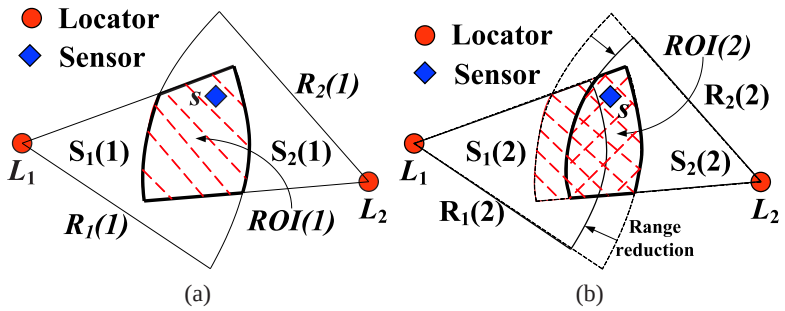


Fig. 7. (a) The sensor is located within the intersection of the sectors $S_1(j), S_2(j)$, which defines the ROI, (b) the locators reduce their communication range and transmit updated beacons. While s is outside the communication range of L_1, it can still hear the transmission of L_2. The new beacon information leads to the reduction of the ROI.

$S_i(R(j), j)$. To illustrate the ROI reduction, consider Figure 7(a), where locators L_1, L_2 transmit with their maximum power; sensor s computes: $ROI(1) = S_1(1) \cap S_2(1)$. In Figure 7(b), locators L_1, L_2 reduce their communication range by lowering their transmission power and re-transmit the updated beacons. While locator L_1 is out of range from sensor s and, hence, does not further refine the sensor's location, s can still hear locator L_2 and therefore, reduce the size of the ROI.

3. Hybrid approach: The combination of the variation of the antenna orientation and communication range leads to a dual dependency of the sector area $S_i(\theta_i(j), R(j), j)$. Such a dependency can also be interpreted as a limited mobility model for the locators. For a locator L_i moving in a confined area, the antenna orientation and communication range with respect to a static sensor varies, thus providing the sensor with multiple sector areas $S_i(j)$. The mobility model is characterized as limited, since the locator has to be within the range of the sensor for at least a fraction of its transmissions in order to provide the necessary localization information. We now present the algorithmic details of HiRLoc.

4.4 Securing the Beacon Transmissions

We now describe the mechanisms used to secure the beacons transmitted by the locators.

Encryption: All beacons transmitted from locators are encrypted with a globally shared symmetric key K_0. Although K_0 can easily be compromised with the compromise of a single sensor, this solution is adopted for resource efficiency reasons. Using K_0, Locators are able to broadcast the localization information, instead of unicasting the information to each sensor. Stronger broadcast authentication algorithms known for ad hoc networks, require the existence of a central authority and time synchronization among all nodes of the network [29]. In Section 5, we show

SeRLoc: Secure Range-Independent Localization Scheme

L : **broadcast** L_i : $\{\ (X_i, Y_i)\ ||\ (\theta_1, \theta_2)\ ||\ (H^{n-j}(PW_i))\ \|\ j\ \|\ ID_{L_i}\ \}_{K_0}$
$LH_s = \{L_i : \|s - L_i\| \leq R\} \bigcap \{H(H^{n-j}(PW_i)) = H^{n-j+1}(PW_i)\}$
s : **define** $A_s = [X_{max} - R,\ X_{min} + R,\ Y_{max} - R,\ Y_{min} + R]$
for k=1:res
  for w=1:res
    $g(k,w) = (x_{g_i}, y_{g_i}) = \left(X_{max} - R + k\frac{X_{max}-X_{min}}{res},\ Y_{max} - R + w\frac{Y_{max}-Y_{min}}{res}\right)$
    for $z = 1 : |LH_s|$
      if $\{\|g(k,w) - L_z\| \leq R\} \bigcap \{\theta_1 \leq \angle g(k,w) \leq \theta_2\}$
        $GST(k,w) = GST(k,w) + 1$
$MG_s = \{g(k,w) : \{k,w\} = \arg\max GST\}$

$$\tilde{s} : (x_{est}, y_{est}) = \left(\frac{1}{|MG_s|} \sum_{i=1}^{|MG_s|} x_{g_i}, \frac{1}{|MG_s|} \sum_{i=1}^{|MG_s|} y_{g_i} \right)$$

Fig. 8. The pseudocode of SeRLoc.

that sensors are able to detect attacks even if K_0 has been compromised, using consistency checks.

In addition to K_0, every sensor s shares a symmetric pairwise key K_{s,L_i} with every locator L_i, also preloaded. Since the number of locators deployed is relatively small, the storage requirement at the sensor side is within the storage constraints (a total of $|L|$ keys). For example, mica motes [24] have 128Kbytes of programmable flash memory. Using 64-bit RC5 [32] symmetric keys and for a network with 400 locators, a total of 3.2Kbytes of memory is required to store all the keys of the sensor with every locator. In order to save storage space at the locator (locators would have to store $|S|$ keys), pairwise keys K_{s,L_i} are derived by a master key K_{L_i}, using a pseudorandom function h [37], and the unique sensor ID_s: $K_{s,L_i} = h_{K_{L_i}}(ID_s)$.

Locator ID Authentication: We use the following scheme based on *efficient one-way hash chains* [15], to provide locator ID authentication. Each locator L_i has a unique password PW_i, blinded with the use of a collision-resistant hash function such as SHA1 [37]. Due to the collision resistance property, it is computationally infeasible for an attacker to find a PW_j, such that $H(PW_i) = H(PW_j),\ PW_i \neq PW_j$. The hash sequence is generated using the following equation:

$$H^0 = PW_i,\ \ H^i = H(H^{i-1}),\ \ i = 1, \cdots, n,$$

with n being a large number and H^0 never revealed to any sensor. Each sensor is preloaded with a table containing the ID of each locator and the corresponding hash value $H^n(PW_i)$. For a network with 400 locators, we need 9 bits to represent locator IDs. In addition, collision-resistant hash functions such as SHA1 [37] have a 160-bit output. Hence, the storage requirement of the hash table at any sen-

HiRLoc: High-resolution Robust Localization Scheme

L : **broadcast** $L_i : \{ (X_i, Y_i) \,||\, (\theta_{i,1}(1), \theta_{i,2}(1)) \,||\, R_i(1)\}$
s : **define** $LH_s = \{L_i : \|s - L_i\| \leq R_i(1)\}$
s : **define** $A_s = [X_{max} - R_i(1),\ X_{min} + R_i(1),\ Y_{max} - R_i(1),\ Y_{min} + R_i(1)]$
s : **store** $S \leftarrow S_i(1) : \{ (X_i, Y_i) \,||\, (\theta_{i,1}(1), \theta_{i,2}(1)) \,||\, R_i(1)\},\ \forall L_i \in LH_s$
$j = 1$
for $k = 1 : Q - 1$
  for $w = 1 : N - 1$
    $j++$
    L **reduce** $R(j) = R(j-1) - \frac{R(1)}{N}$
    L : **broadcast** $L_i : \{ (X_i, Y_i) \,||\, (\theta_{i,1}(j), \theta_{i,2}(j)) \,||\, R_i(j)\}$
    s : **replace** $S \leftarrow S_i(j) : \{ (X_i, Y_i) \,||\, (\theta_{i,1}(j), \theta_{i,2}(j)) \,||\, R_i(j)\},$
      $\forall L_i : \|s - L_i\| \leq R_i(j) \bigcap L_i \in LH_s$
  endfor
  $j++$
  $R_i(j) = R_i(1),\ \forall L_i \in LH_s$
  L **rotate** $\theta_i(j) = \{\theta_{i,1}(j-1) + \frac{2\pi}{MQ}, \theta_{i,2}(j-1) + \frac{2\pi}{MQ}\}$
  L : **broadcast** $L_i : \{ (X_i, Y_i) \,||\, (\theta_{i,1}(j), \theta_{i,2}(j)) \,||\, R_i(j)\}$
  s : **store** $S \leftarrow S_i(j) : \{ (X_i, Y_i) \,||\, (\theta_1(j), \theta_2(j)) \,||\, R_i(j)\},\ \forall L_i : \|s - L_i\| \leq R(j) \bigcap L_i \in LH_s$
endfor
s : **compute** $ROI = \bigcap_{i=1}^{|S|} S_i$

Fig. 9. The pseudocode of HiRLoc.

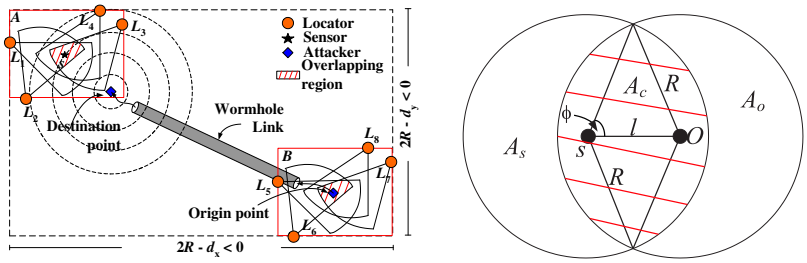
sor is 8.45Kbytes[2]. To reduce the storage needed at the locators, we employ an efficient storage/computation method for hash chains of time/storage complexity $\mathcal{O}(\log^2(n))$ [8].

The j^{th} broadcasted beacon from locator L_i includes the hash value $H^{n-j}(PW_i)$, along with the index j. Every sensor that hears the beacon accepts the message only if $H(H^{n-j+1}(PW_i)) = H^{n-j}(PW_i)$. After verification, the sensor replaces $H^{n-j+1}(PW_i)$ with $H^{n-j}(PW_i)$ in its memory and increases the hash counter by one so as to perform only one hash operation in the reception of the next beacon from the same locator L_i. The index j is included in the beacons so that sensors can resynchronize with the current published hash value in case of loss of some intermediate hash values. The beacon of locator L_i has the following format:

$$L_i : \{ (X_i, Y_i) \,||\, (\theta_1, \theta_2) \,||\, (H^{n-j}(PW_i)) \,\|\, j \,\|\, ID_{L_i} \}_{K_0},$$

where $||$ denotes the concatenation operation and $\{m\}_K$ denotes the encryption of message m with key K. Note that our method does not provide end-to-end locator authentication, but only guarantees authenticity for the messages received from locators directly heard to a sensor. This condition is sufficient to secure our localization

[2] The required storage at each sensor in order to store 400 64-bit RC5 keys, 400 160-bit SHA1 hash values for secure communication with 400 locators is now 11.65Kbytes.



(a) (b)

Fig. 10. (a) Wormhole attack: an attacker records beacons in area B, tunnels them via the wormhole link in area A and rebroadcasts them. (b) Computation of the common area A_c, where locators are heard to both s, O.

scheme against possible attacks. The pseudocode for SeRLoc is presented in Figure 8. The pseudocode for HiRLoc is presented in Figure 9.

5 Threat Analysis

In this section, we show how SeRLoc and HiRLoc are resilient to the attacks described in Section 2. Note that our goal to allow sensors to determine their location, even in the presence of attacks and not to prevent attacks that may be harmful in other network protocols.

5.1 The Wormhole Attack

Threat Model

In the case of our location estimation process an attacker launching a wormhole attack records the beacons transmitted from locators at the origin point of the attack and replays them at the destination point, thus providing false localization information to the sensors attacked. In Figure 10(a), the attacker records beacons at region B, tunnels them via the wormhole link in region A, and replays them, thus leading sensor s to believe that it can hear locators $\{L_1 \sim L_8\}$.

Detecting Wormholes

In the case of a wormhole attack, the cryptography used to secure the beacon transmissions, and to authenticate the source of the information is not violated. Wormholes violate the geometry of the network by enabling the propagation of messages at a distance longer than the communication range [30]. Hence, in the case of the wormhole attack, additional non-cryptographic mechanisms are needed to detect the geometry violation. We now show how a sensor can detect a wormhole attack using two consistency check properties: the *single message/sector per locator* property and

the *communication range constraint* property.

Single Message/Sector per Locator Property: The origin point O of the wormhole attack defines the set of locators LH_s^r replayed to the sensor s under attack. The location of the sensor defines the set of locators LH_s^d directly heard to the sensor s, with $LH_s = LH_s^r \cup LH_s^d$. Based on the single message/sector per locator property we show that the wormhole attack is detected when $LH_s^r \cap LH_s^d \neq \emptyset$.

Lemma 1. *Single message per locator/sector property: reception of multiple messages authenticated with the same hash value is due to replay, multipath effects, or imperfect sectorization.*

Proof. In the absence of any attack, a sensor can hear multiple sectors due to multipath effects. In addition, a sensor located at the boundary of two sectors can also hear multiple sectors even if there is no multipath or attack. We assume that the same but fresh hash value is used to authenticate them per beacon transmission. Hence, sensors will only accept the first message arriving from any sector of the same locator, per transmission. Due to the use of an identical but fresh hash in all sectors per transmission, if an adversary replays a message from any sector of a locator directly heard by the sensor under attack, the sensor will have already received the hash via the direct path and, hence, detect the attack and reject the message.

If we consider reception of multiple messages containing the same hash value due to multipath effects or imperfect sectorization to be a replay attack, a sensor will always assume it is under attack when it receives messages with the same hash value. Hence, an adversary launching a wormhole attack will always be detected if it replays a message from locator $L_i \in LH_s^d$, that is, if $LH_s^r \cap LH_s^d \neq \emptyset$. In Figure 11(a), A_s denotes the area where, $L_i \in LH_s^d$ (circle of radius R centered at s), A_o denotes the area where $L_i \in LH_s^r$ (circle of radius R centered at O), and the shaded area A_c denotes the common area $A_c = A_s \cap A_o$.

Claim. The detection probability $P(SG)$ due to the single message/sector per locator property is equal to the probability that at least one locator lies within an area of size A_c, and is given by:

$$P(SG) = 1 - e^{-\rho_L A_c}, \quad \text{with } A_c = 2R^2\phi - Rl\sin\phi, \quad \phi = \cos^{-1}\frac{l}{2R}. \quad (10)$$

with l as the distance between the origin point and the sensor under attack.

Proof. If a locator L_i lies inside A_c, it is less than R units away from a sensor s and, therefore $L_i \in LH_s^d$. Locator L_i is also less than R units away from the origin point of the attack O, and therefore, $L_i \in LH_s^r$. Hence, if a locator lies inside A_c, $LH_s^r \cap LH_s^d \neq \emptyset$, and the attack is detected due to the single message/sector per locator property. The detection probability $P(SG)$ is equal to the probability that at least one locator lies within A_c. If LH_{A_c} denotes the set of locators located within area A_c then:

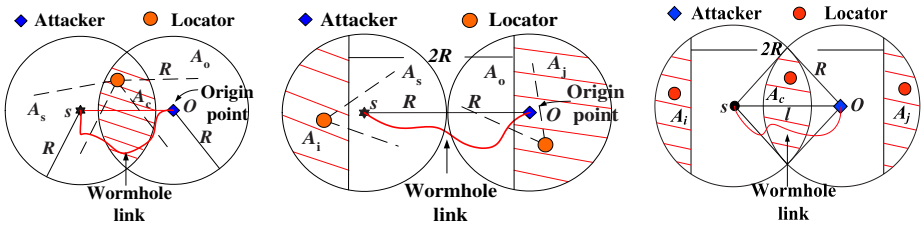


Fig. 11. (a) Single message/sector per locator property: a sensor s cannot hear two messages authenticated with the same hash value. (b) Communication range violation property: a sensor s cannot hear two locators more than $2R$ apart. (c) Combination of the two properties for wormhole detection.

$$P(SG) = P(|LH_{A_c}| \geq 1) = 1 - P(|LH_{A_c}| = 0) = 1 - e^{-\rho_L A_c}, \qquad (11)$$

where A_c can be computed from Figure 10(b) to be:

$$A_c = 2R^2\phi - Rl\sin\phi, \quad \phi = \cos^{-1}\frac{l}{2R}, \qquad (12)$$

with $l = \|s - O\|$.

Figure 12(a) presents the detection probability $P(SG)$ vs. the locator density ρ_L and the distance $\|s - O\|$ between the origin point and the sensor under attack, normalized over R. We observe that if $\|s - O\| \geq 2R$, then $A_c = 0$, and the use of the single message/sector per locator property is not sufficient to detect a wormhole attack. For distances $\|s - O\| \geq 2R$, a wormhole attack can be detected using the following communication range constraint property.

Communication Range Violation Property: Given the coordinates of node s, all locators LH_s heard by s should lie within a circle of radius R, centered at s. Since node s is not aware of its location it relies on its knowledge of the locator-to-sensor communication range R to verify that the set LH_s satisfies Lemma 2.

Lemma 2. *Communication range constraint property: A sensor s cannot hear two locators $L_i, L_j \in LH_s$, more than $2R$ apart, that is, $\|L_i - L_j\| \leq 2R$, $\forall L_i, L_j \in LH_s$.*

Proof. Any locator $L_i \in LH_s$ has to lie within a circle of radius R, centered at the sensor s (area A_s in Figure 11(b)), $\|L_i - s\| \leq R, \forall L_i \in LH_s$. Hence,

$$\|L_i - L_j\| = \|L_i - s + s - L_j\| \leq \|L_i - s\| + \|s - L_j\| \leq R + R = 2R. \qquad (13)$$

Using the coordinates of LH_s, a sensor can detect a wormhole attack if the communication range constraint property is violated. We now compute the detection probability $P(CR)$ due to the communication range constraint property.

Claim. A wormhole attack is detected due to the communication range constraint property, with a probability:

$$P(CR) \geq \left(1 - e^{-\rho_L A_i^*}\right)^2, \quad A_i^* = x\sqrt{R^2 - x^2} - R^2 \tan^{-1}\left(\frac{x\sqrt{R^2 - x^2}}{x^2 - R^2}\right) \quad (14)$$

where $x = \frac{\|s-O\|}{2}$.

Proof. Consider Figure 11(b), where $\|s - O\| = 2R$. If any two locators within A_s, A_o have a distance larger that $2R$, a wormhole attack is detected. Though $P(CR)$ is not easily computed analytically, we can obtain a lower bound on $P(CR)$ by considering the following event. In Figure 11(b), the vertical lines defining shaded areas A_i, A_j, are perpendicular to the line connecting s, O, and have a separation of $2R$. If there is at least one locator L_i in the shaded area A_i and at least one locator L_j in the shaded area A_j, then $\|L_i - L_j\| \rangle 2R$ and the attack is detected. Note that this event does not include all possible locations of locators for which $\|L_i - L_j\| \rangle 2R$, and hence it yields a lower bound. If $\mathcal{LH}_{A_i,A_j}$ denotes the event $\left(|LH_{A_i}| \rangle 0 \cap |LH_{A_j}| \rangle 0\right)$ then,

$$\begin{aligned}
P(CR) &= P(\|L_i - L_j\| \rangle 2R, L_i, L_j \in LH_s) \\
&\geq P(CR \bigcap \mathcal{LH}_{A_i,A_j}) && (15) \\
&= P\left(CR \mid \mathcal{LH}_{A_i,A_j}\right) P(\mathcal{LH}_{A_i,A_j}) && (16) \\
&= P(\mathcal{LH}_{A_i,A_j}) && (17) \\
&= (1 - e^{-\rho_L A_i})(1 - e^{-\rho_L A_j}), && (18)
\end{aligned}$$

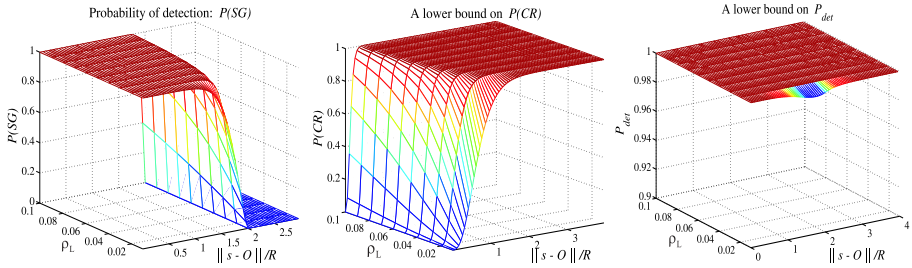
where (15) follows from the fact that the probability of the intersection of two events is always less or equal to the probability of one of the events, (16) follows from the definition of the conditional probability, (17) follows from the fact that when $\mathcal{LH}_{A_i,A_j}$ is true, we always have a communication range constraint violation ($P(CR \mid \mathcal{LH}_{A_i,A_j}) = 1$), and (18) follows from the fact that A_i, A_j are disjoint areas and that locators are randomly deployed.

We can maximize the lower bound of $P(CR)$, by finding the optimal values A_i^*, A_j^*. In fact it can be shown that the lower bound in (18) attains its maximum value when $A_i^* = \max_i\{A_i\}$ subject to the constraint $A_i = A_j$ (A_i, A_j are symmetric) [17]. and is given by:

$$A_i^* = A_j^* = x\sqrt{R^2 - x^2} - R^2 \tan^{-1}\left(\frac{x\sqrt{R^2 - x^2}}{x^2 - R^2}\right), \quad \text{and } x = \frac{\|s - O\|}{2}. \quad (19)$$

Inserting (19) into (18) yields the required result: $P(CR) \geq \left(1 - e^{-\rho_L A_i^*}\right)^2$.

In Figure 12(b), we show the maximum lower bound on $P(CR)$ vs. the locator density ρ_L, and the distance $\|s - O\|$ normalized over R. The lower bound on



(a) (b) (c)

Fig. 12. Wormhole detection probability based on, (a) the single message/sector per locator property: $P(SG)$. (b) A lower bound on the wormhole detection based on the communication range violation property: $P(CR)$. (c) A lower bound on the wormhole detection probability for SeRLoc.

$P(CR)$ increases with the increase of $\|s-O\|$ and attains its maximum value for $\|s-O\|=4R$ when $A_i^*=A_j^*=\pi R^2$. For distances $\|s-O\|\rangle 4R$ a wormhole attack is always detected based on the communication range constraint property, since any locator within A_o will be more than $2R$ apart from any locator within A_s.

Detection Probability P_{det} of the Wormhole Attack: We now combine the two detection mechanisms, namely the single message/sector per locator property and the communication range constraint property for computing the detection probability of a wormhole attack.

Claim. The detection probability of a wormhole attack is lower bounded by $P_{det} \geq (1-e^{-\rho_L A_c})+(1-e^{-\rho_L A_i^*})^2 e^{-\rho_L A_c}$.

Proof. In the computation of the communication range constraint property, by setting $A_i = A_j$ and maximizing A_i regardless of the distance $\|s-O\|$, the areas A_i, A_j, and A_c do not overlap as shown in Figure 11(c). Hence, the corresponding events of finding a locator at any of these areas are independent and we can derive a lower bound on the detection probability P_{det} by combining the two properties.

$$\begin{aligned} P_{det} = P(SG \cup CR) &= P(SG)+P(CR)-P(SG)P(CR) \\ &= P(SG)+P(CR)\,(1-P(SG)) \\ &\geq (1-e^{-\rho_L A_c})+(1-e^{-\rho_L A_i^*})^2 e^{-\rho_L A_c}. \end{aligned} \tag{20}$$

The left side of (20) is a lower bound on P_{det} since $P(CR)$ was also lower bounded.

In Figure 12(c), we show the lower bound on P_{det} vs. the locator density ρ_L and the distance $\|s-O\|$ normalized over R. For values of $\|s-O\|\rangle 4R$, $P_{CR}=1$, since any $L_i \in LH_s^d$ will be more than $2R$ away from any $L_j \in LH_s^r$ and hence, the wormhole attack is always detected. From Figure 12(c), we observe that a wormhole attack is detected with a probability very close to unity, independent of the origin and destination point of the attack.

Attach to Closer Locator Algorithm (ACLA)

s : **broadcast** $\{ \eta_s \parallel ID_s \}$
if L_i *hears* $\{ \eta_s \parallel ID_s \}$ **reply**
$L_i : \{ \eta_s \parallel (X_i, Y_i) \,||\, (\theta_1, \theta_2) \,||\, (H^{n-j}(PW_i)) \parallel j \parallel ID_{L_i} \}_{K_{s,L_i}}$
L'_i : *first authentic reply from a locator.*
$LH_s^d = \{L_i \in LH_s$: *sector*$\{L_i\}$ *intersects sector*$\{L'_i\}\}$
s : **execute** *SeRLoc with* $LH_s = LH_s^d$

Fig. 13. The pseudocode of ACLA.

Location Resolution Algorithm: Although a wormhole can be detected using one of the two detection mechanisms, a sensor s under attack cannot distinguish the set of locators directly heard LH_s^d from the set of locators replayed LH_s^r and hence, estimate its location. To resolve the location ambiguity sensor s executes the *Attach to Closer Locator Algorithm* (ACLA).

Assume that a sensor authenticates a set of locators $LH_s = LH_s^d \cup LH_s^r$, but detects that it is under attack.

- *Step 1:* Sensor s broadcasts a randomly generated nonce η_s and its ID_s.
- *Step 2:* Every locator hearing the broadcast of sensor s replies with a beacon that includes localization information and the nonce η_s, encrypted with the pairwise key K_{s,L_i} instead of the broadcast key K_0. The sensor identifies the locator L'_i that replies first with an authentic message that includes η_s.
- *Step 3:* Sensor s identifies the set LH_s^d as all the locators whose sectors overlap with the sector of L'_i, and executes SeRLoc with $LH_s = LH_s^d$.

The pseudocode of ACLA is presented in Figure 13. Note that the closest locator to sensor s will always reply first if it directly hears the broadcast from s, and not through a replay from an adversary. In order for an adversary to force sensor s to accept set LH_s^r as the valid locator set, it can only replay the nonce η_s to a locator $L_i \in LH_s^r$, record the reply, tunnel via the wormhole and replay it in the vicinity of s. However, a reply from a locator in LH_s^r will arrive later than any reply from a locator in LH_s^d, since locators in LH_s^r are further away from s than locators in LH_s^d.

To execute ACLA, a sensor must be able to communicate bidirectionally with at least one locator. The probability $P_{s\to L}$ of a sensor having a bidirectional link with at least one locator and the probability P_{bd} that *all* sensors can bidirectionally communicate with at least one locator can be computed as:

$$P_{s\to L} = 1 - e^{-\rho_L \pi r^2 G^{\frac{2}{\gamma}}}, \quad P_{bd} = (1 - e^{-\rho_L \pi r^2 G^{\frac{2}{\gamma}}})^{|S|}. \tag{21}$$

Hence, we can select the system parameters ρ_L, G so every sensor has a bidirectional link with at least one locator with any desired probability.

5.2 Impersonation (Sybil) Attack

An adversary can launch an impersonation attack against SeRLoc or HiRLoc if it successfully impersonates locators. Since sensors are pre-loaded with valid locator IDs along with the hash values corresponding to the head of the reversed hash chain for each locator, an adversary can only impersonate locators by compromising the globally shared key K_0.

Once K_0 has been compromised, the adversary has access to both locators IDs, the hash chain values published by the locators, as well as the coordinates of the locators. Since sensors always have the latest published hash values from the locators that they directly hear, an adversary can only impersonate locators that are not directly heard to the sensors under attack. The adversary can generate bogus beacons, attach an already published hash value from a locator not heard by the sensor under attack, and encrypt it with the compromised K_0.

Depending on the type of locators used, static or mobile, an adversary can impersonate locators in different ways. If the locators are static and their location is known before deployment, the coordinates of all locators can be preloaded to every sensor. Hence, the adversary cannot advertise a location that is different from the actual coordinates of an impersonated locator. In such a case, the Sybil attack is equivalent to a replay attack since the adversary cannot alter the content of the beacons[3]. If the locators are mobile, or their coordinates cannot be preloaded to the sensors before deployment, the adversary can place the impersonated locators to arbitrary positions. Hence, by impersonating a higher number of locators than the ones directly heard by the sensor under attack, the adversary can compromise the majority vote scheme of SeRLoc and displace the sensor.

Defense against the Sybil Attack: Though we do not provide a mechanism to prevent an adversary from impersonating locators except for the ones directly heard by a sensor, we can still determine the position of sensors in the presence of Sybil attack. In the case where sensors know a priori the coordinates of the locators, the sensor can detect the Sybil attack with the same mechanisms used for the wormhole attack, since the Sybil attack becomes a beacon replay. In the case where the coordinates of the locators are not preloaded to the sensors, an adversary can manipulate the coordinates of the impersonated locators, so that neither of the wormhole defense mechanisms detect an anomaly. The adversary needs to impersonate more than LH_s^d locators in order to displace the sensor s. To avoid sensor displacement we rely on the invariability of the locator deployment statistics to detect locator impersonation.

Since the locator density ρ_L is known before deployment, we can select a threshold value L_{max} as the maximum allowable number of locators heard by each sensor. If a sensor hears more than L_{max} locators, it assumes that it is under attack and executes ACLA to determine its position. The probability that a sensor s hears more than L_{max} locators is given by:

[3] The adversary can alter the angle information contained in the beacon. However, this is equivalent to replaying the beacon of another sector.

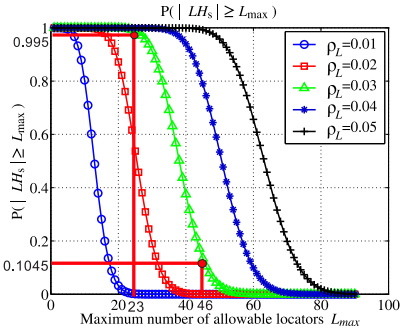


Fig. 14. $P(|LH_s| \geq L_{max})$, vs. L_{max} for varying locator densities ρ_L. When $\rho_L = 0.03$, a choice of $L_{max} = 46$ allows a sensor to localize itself when under Sybil attack with a probability $P(|LH_s| \geq 23) = 0.995$, while the false positive alarm probability is $P(|LH_s| \rangle 46) = 0.1045$.

$$P(|LH_s| \geq L_{max}) = 1 - P(|LH_s| \langle L_{max}))$$
$$= 1 - \sum_{i=0}^{L_{max}-1} \frac{(\rho_L \pi R^2)^i}{i!} e^{-\rho_L \pi R^2}. \quad (22)$$

Using (22), we can select the value of L_{max} so that there is a very small probability for a sensor to hear more than L_{max} locators, while there is a very high probability for a sensor to hear more than $\frac{L_{max}}{2}$ locators. If a sensor hears more than L_{max} locators without being under attack, the detection mechanism will result in a false positive alarm and force the sensor to execute ACLA to successfully locate itself. However, if a sensor hears less than $\frac{L_{max}}{2}$, the sensor is vulnerable to a Sybil attack. Therefore, we must select a threshold L_{max} so that any sensor hears less than $\frac{L_{max}}{2}$ locators with a probability very close to zero.

In Figure 14, we show $P(|LH_s| \geq L_{max})$ vs. L_{max}, for varying locator densities ρ_L. Based on Figure 14, we can select the appropriate L_{max} for each value of ρ_L. For example, when $\rho_L = 0.03$, a choice of $L_{max} = 46$ allows a sensor to localize itself when under Sybil attack with a probability $P(|LH_s| \geq 23) = 0.995$, while the false positive alarm probability is $P(|LH_s| \rangle 46) = 0.1045$.

5.3 Compromised Network Entities

In this section, we examine the robustness of SeRLoc and HiRLoc to compromised network entities. We consider a sensor node or a locator node to be compromised if an attacker assumes full control over the behavior of the node and knows all the keys stored at the compromised node.

Compromised Sensors: Though sensors are assumed to be easier to compromise, an attacker has no incentive to compromise sensors, since they do not actively participate in the localization procedure. The only benefit in compromising a sensor is

to gain access to the globally shared key K_0.

Compromised Locators: An adversary that compromises a locator L_i gains access to the globally shared key K_0, the pairwise keys K_{s,L_i} shared between the locator and every sensor, as well as all the hash values of the locator's hash chain. By compromising a single locator, the adversary can displace any sensor, by impersonating the compromised locator from a position closer to the sensor under attack compared to the closest legitimate locator. The adversary impersonates multiple locators in order to force location ambiguity to the sensor under attack. Once the attack is detected, sensor s executes ACLA to resolve its location ambiguity. Since the adversary is closer to the sensor s than the closest legitimate locator, its reply will arrive to s first. Hence, s will assume that the impersonated set of locators is the valid one and will be displaced.

To avoid sensor displacement by a single locator compromise, we can intensify the resilience to locator compromise by involving more than one locators in the location resolution algorithm at the expense of higher communication overhead. A sensor s under attack, can execute the *Enhanced Location Resolution Algorithm* (ELRA) that follows.

- *Step 1:* Sensor s broadcasts a randomly generated nonce η_s, the set of locators heard LH_s and its ID_s.

$$s: \{ \eta_s \parallel LH_s \parallel ID_s \}. \tag{23}$$

- *Step 2:* Every locator L_i receiving the broadcast from s appends its coordinates, the next hash value of its hash chain and its ID_{L_i}, encrypts the message with K_0 and re-broadcasts the message to all sectors.

$$L_i: \{\eta_s \parallel LH_s \parallel ID_s \parallel (X_i, Y_i) \parallel H^{n-k}(PW_i) \parallel \parallel j \parallel ID_{L_i} \}_{K_0}. \tag{24}$$

- *Step 3:* Every locator receiving the rebroadcast, verifies the authenticity of the message, and that the transmitting locator is within its range. If the verification is correct and the receiving locator belongs to LH_s, the locator broadcasts a new beacon with location information and the nonce η_s encrypted with the pairwise key K_{s,L_i} with sensor s.

$$L_i: \{ \eta_s \parallel (X_i, Y_i) \parallel (\theta_1, \theta_2) \parallel H^{n-k}(PW_i) \parallel j \parallel ID_{L_i} \}_{K_{s,L_i}}. \tag{25}$$

- *Step 4:* The sensor collects the first L_{max} authentic replies from locators and executes SeRLoc with $LH_s = L_{max}$.

The pseudocode for the enhanced location resolution algorithm is presented in Figure 15. Note that for a locator to hear the sensor's broadcast, it has to be within a range $r_{sL} = rG^{\frac{1}{\gamma}}$ from the sensor. Furthermore, in order for a the sensor to make the correct location estimate, all locators within a range R from s need to provide new beacon information.

Claim. Every locator positioned within R from a sensor s is within the range of any locator positioned at a distance r_{sL} from the sensor s.

Enhanced Location Resolution Algorithm (ELRA)

s : **broadcast** $\{ \eta_s \parallel LH_s \parallel ID_s \}$
$RL_s = \{L_i : \|s - L_i\| \leq r_{sL}\}$
RL_s : **broadcast** $\{ \eta_s \parallel LH_s \parallel ID_s \parallel (X_i, Y_i) \parallel H^{n-k}(PW_i) \parallel j \parallel ID_{L_i} \}_{K_0}$ $BL_s = \{L_i : \|RL_s - L_i\| \leq r_{LL}\} \bigcap LH_s$
BL_s : **broadcast** $\{ \eta_s \parallel (X_i, Y_i) \parallel (\theta_1, \theta_2) \parallel H^{n-k}(PW_i) \parallel j \parallel ID_{L_i} \}_{K_{s,L_i}}$
s : **collect** *first* L_{max} *authentic beacons from* BL_s
s : **execute** *SeRLoc with collected beacons*

Fig. 15. The pseudocode for the enhanced location resolution algorithm (ELRA).

Proof. For any locator positioned at a distance r_{sL} from the sensor s to reach any locator positioned at a distance R from sensor s, the following condition has to hold: $r_{LL} \geq R + r_{sL}$.

$$RG^{\frac{2}{\gamma}} \geq R + rG^{\frac{1}{\gamma}} \Rightarrow \frac{R}{rG^{\frac{1}{\gamma}}}(G^{\frac{2}{\gamma}} - 1) \geq 1. \tag{26}$$

Since $R \geq rG^{\frac{2}{\gamma}}$ by assumption, and $G^{\frac{2}{\gamma}} \geq 1$, the left side of (26) is always greater than one.

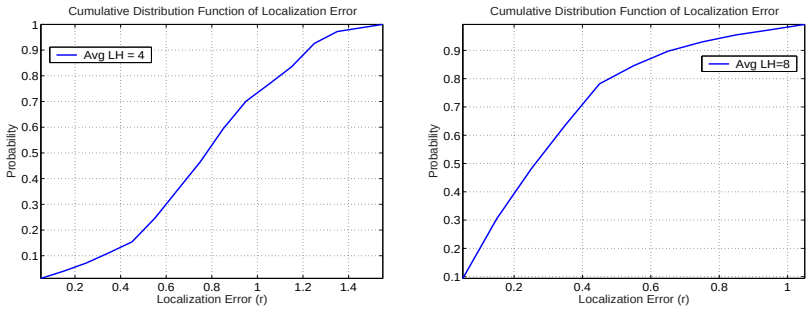
Each beacon broadcast from a locator has to include the nonce η_s initially broadcasted by the sensor and be encrypted with the pairwise key between the sensor and the locator. Hence, given that the sensor has at least $\frac{L_{max}}{2}$ locators within range R with very high probability (see Figure 14), the adversary has to compromise at least $\left(\frac{L_{max}}{2} + 1\right)$ locators, in order to compromise the majority vote scheme of SeRLoc. In addition, the attacker has to possess the hardware capabilities to process and transmit $\left(\frac{L_{max}}{2} + 1\right)$ replies before $\frac{L_{max}}{2}$ replies from valid locators reach the sensor under attack. Our enhanced location resolution algorithm significantly increases the resilience of SeRLoc to locator compromise at the expense of higher communication overhead at the locators.

6 Performance Evaluation

In this section, we evaluated the performance of SeRLoc and HiRLoc with respect to their localization accuracy. To emulate the conditions of a real deployment, we also evaluated SeRLoc under error in the locators' coordinates and false estimation of the antenna sector that includes the sensors and empirically showed that SeRLoc is robust against both sources of error.

6.1 Simulation Setup

We randomly distributed 5,000 sensors within a $100\text{x}100m^2$ rectangular area. We also randomly placed locators within the same area and computed the average localization error as:



(a) (b)

Fig. 16. The cumulative distribution function (cdf) of the localization error of SeRLoc when $M = 3$ and, (a) $\overline{LE} = 4$, (b) $\overline{LH} = 8$.

$$\overline{LE} = \frac{1}{|S|} \sum_i \frac{\|\tilde{s}_i - s_i\|}{r}, \tag{27}$$

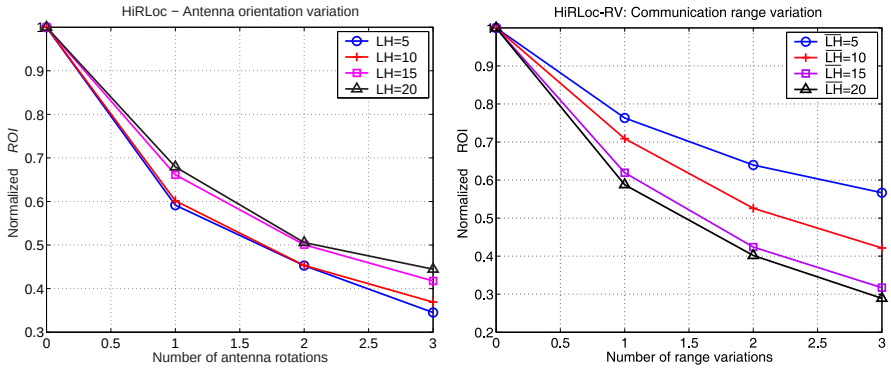
where S is the set of sensors, $\tilde{s}_i$ is the sensor estimated position, s_i is the real position and r is the sensor-to-sensor communication range.

6.2 Localization Error vs. Locators Heard

In our first experiment, we investigated the impact of the average number of locators heard $\overline{LH}$ on the localization error. In Figures 16(a) and (b), we show the cumulative distribution function (cdf) of the localization error for SeRLoc when 3-sector antennas are used at the locators and the average number of locators heard are $\overline{LH} = 6$ and $\overline{LH} = 8$, respectively. We observe that for $\overline{LH} = 4$, the error is more evenly distributed among its possible values with 90% of the sensors having an error of less than 1.2r, while for $\overline{LH} = 8$, more than 90% of the sensors have an error smaller than 0.7r.

The highest localization error occurs when a sensor hears only one locator L_i and is R units away from L_i. The probability for such an event to occur can be set to an arbitrary small value by deploying a sufficient number of locators. For example, when $\overline{LH} = 8$, the probability for a sensor to hear just one locator is $P(|LH| = 1) = 2.7\text{x}10^{-3}$.

In Figure 17(a) we show the ROI vs. the number of antenna rotations, and for varying $\overline{LH}$, when 3-sector antennas are used at each locator. Note that the ROI is normalized over the size of the ROI given by SeRLoc denoted by ROI(1) (no antenna rotation). From Figure 17(a), we observe that even a single antenna rotation, reduces the size of the ROI by more than 50%, while three antenna rotations reduce the size to $ROI(4) = 0.12ROI(1)$, when $\overline{LH} = 5$. A reduction of 50% in the size of the ROI by a single antenna rotation means that one can deploy half the locators compared to SeRLoc and achieve the same localization accuracy by just rotating the locators' antennas once. The savings in locators are significant considering that the reduction in hardware requirements comes at no additional communication cost.



(a) (b)

Fig. 17. Normalized ROI vs. number of antenna rotations for varying $\overline{LH}$. The ROI is normalized with respect to the ROI acquired with no variation of the antenna orientation (SeRLoc). (b) ROI vs. number of range reductions for varying $\overline{LH}$.

We also observe that as $\overline{LH}$ increases, HiRLoc provides diminishing returns. This is due to the fact that when the number of locators heard at each sensor is high, SeRLoc already provides a good location estimate (small ROI) and, hence, the margin for reduction of the ROI size is limited. In Figure 17(b) we show the normalized ROI vs. the number of communication range reductions, and for different $\overline{LH}$ values, when locators are equipped with 3-sector antennas.

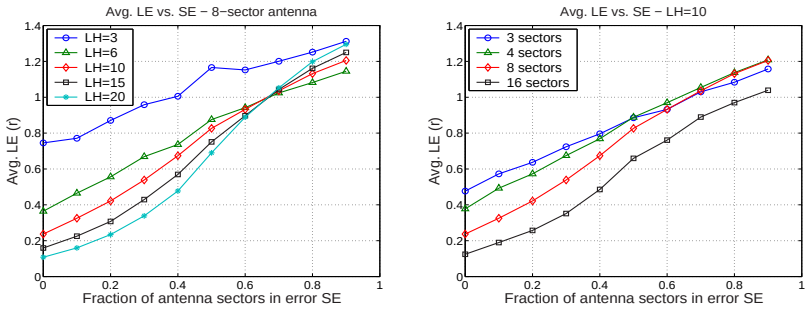
From Figure 17(b), we observe that the communication range variation, though significantly improves the system performance, does not achieve the same ROI reduction as the antenna orientation variation[4]. This behavior is explained by the fact that the gradual reduction of the communication range reduces the number of beacons heard at each sensor, in contrast with the antenna orientation variation case where the same number of locators is heard at the sensors at each antenna rotation. In addition, we observe that greater ROI reduction occurs when the $\overline{LH}$ at each locator is high. This is justified by considering that a higher $\overline{LH}$ allows for more sectors with lower communication range to intersect and hence, smaller ROI.

6.3 Localization Error vs. Sector Error

Sensors may be located close to the boundary of two sectors of a locator, or be deployed in a region with high multipath effects. In such a case, a sensor may falsely assume that it is located in another sector, than the actual sector that includes it. We refer to this error as sector error (SE) defined as:

$$SE = \frac{\#\text{ of sectors falsely estimated}}{LH}. \tag{28}$$

[4] The comparison is valid for the same number of $\overline{LH}$, the same number of antenna sectors and the same number of variations in the antenna rotation and communication range, respectively.



(a) (b)

Fig. 18. (a) $\overline{LE}$ vs. sector error SE for varying $\overline{LH}$. (b) Average localization error $\overline{LE}$ vs. sector error SE for varying number of antenna sectors for a network of $|S| = 5,000$ and $\frac{R}{r} = 10$.

A sector error of 0.5 indicates that *every* sensor falsely estimated the sectors of half the locators heard. In Figure 18(a), we show the $\overline{LE}$ vs. the SE for varying $\overline{LH}$, and 8-sector antennas. We observe that the $\overline{LE}$ does not grow significantly large (larger than the sensor communication range r), until a fraction of 0.7 of the sectors are falsely estimated.

SeRLoc is resilient to sector error due to the majority vote scheme employed in the determination of the overlapping region. Even if a significant fraction of sectors are falsely estimated, these sectors do not overlap in the same network area and hence a score low in the grid-sector table.

Note that for a $SE \rangle 0.7$, $\overline{LE}$ increases with $\overline{LH}$. When the SE grows beyond a threshold, the falsely estimated sectors dominate in the location determination. As $\overline{LH}$ grows, the falsely estimated overlapping region, shrinks due to the higher number of overlapping sectors. Therefore, the CoG that defines the sensor's location gets further apart than the actual sensor location.

In Figure 18(b), we show the $\overline{LE}$ vs. SE for $\overline{LH} = 10$ and varying number of antenna sectors. We observe that the narrower the antenna sector the smaller the $\overline{LE}$, even in the presence of SE. For a small SE the overlapping region is dominated by the correctly estimated sectors and shrinks with increasing antenna sectors. For large SE the overlapping region is dominated by the false sectors and an increase in $\overline{LH}$ does not reduce the $\overline{LE}$.

7 Related Work

7.1 Related Work

An extensive literature exists for location estimation schemes for WSN in a benign environment [4, 10, 12, 25, 27, 31, 34–36]. Recently, a number of articles have appeared addressing the problem of sensor location estimation and verification in an adversarial setting [3, 5, 7, 14, 17–22, 33].

Sastry et al. [33] proposed the $ECHO$ protocol for verifying the location claim of a node, using a challenge response scheme and a combination of RF and Ultrasound signals. $ECHO$ is based on a distance bounding protocol proposed by Brands and Chaum [3]. Čapkun and Hubaux proposed Verifiable Multilateration (VM) for securing range-based localization schemes [5]. In VM, a node must verify its distance to at least three reference points in order to securely estimate its position. Čapkun et al. also proposed a location verification method based on hidden reference points that can verify the validity of the location claims of nodes [7].

Liu et al. [23] proposed an attack-resistant location estimation technique that can filter bogus beacon information provided that the majority of significant majority of beacons is benign. Li et al. [21] discuss a variety of attacks specific to the localization process and propose robust statistical methods that provide attack resistant localization. Finally, Kuhn [14] has proposed an asymmetric security mechanism for securing GPS-like navigation signals.

7.2 Open Problems

While the schemes that have been proposed for secure location estimation in WSNs [5, 7, 17–22, 33] are a significant step forward in providing a transparent and secure localization service, several problems remain open. The dependency of the location estimation schemes to physical characteristics such as received signal strength [1], time of arrival or time difference of arrival [27, 34], allows side-channel attacks not related to the strength of the cryptographic primitives used to secure the communication [19, 21, 22].

To combat side-channel attacks a series of consistency checks have been proposed [17–19, 22]. It remains an open problem which of the modalities of a sensor network used to detect attacks against the localization process are invariant to side-channel attacks. The ability of an adversary to alter the physical properties used for localization and distort the environment can significantly impact the localization accuracy.

Furthermore, current secure location estimation techniques do not provide any guarantee on the localization accuracy. The analytical evaluation of the localization error in the presence of adversaries is a problem requiring further investigation. Finally, most secure localization schemes studied localization for static sensor networks. Securing the location estimation process when the reference points, the sensors or both are mobile remains an open problem.

8 Conclusion

In this chapter, we have studied the problem of location estimation for WSN in an adversarial environment. We have demonstrated a series of attacks relevant to range-independent localization methods, such as the relay attack, the impersonation attack and compromise of reference points. We showed that securing the location estimation process requires not only securing the communication link between the

reference points and the sensors, but also additional non-cryptographic consistency checks based on invariant properties such as the communication range or the network deployment statistics.

We proposed a range-independent, decentralized localization scheme called SeRLoc that allows sensors to determine their location in an untrusted environment. We also proposed HiRLoc, a secure location estimation algorithm that relies on the superposition of location information over time to improve the location estimation accuracy. We analytically evaluated the probability of sensor displacement due to security threats in WSNs such as the wormhole attack, the Sybil attack, and compromise of network entities and showed that SeRLoc and HiRLoc provide accurate location estimation even in the presence of these threats. In doing so, we used the geometric and radio range information to detect the attacks on the localization.

Our performance evaluation studies showed that our algorithm are resilient to sources of error such as location error of reference points as well as error in the sector determination. We identified the integration of new modalities for consistency checks, the analytical evaluation of the location estimation error in the presence of adversaries and the secure location estimation for mobile sensor networks as areas of future research.

Acknowledgements

This work was supported in part by the following grants: CAREER grant from NSF ANI-0093187; ARO grant W911NF-05-1-0491; Collaborative Technology Alliance (CTA) from ARL DAAD19-01-2-0011[5].

References

1. P. Bahl and V. Padmanabhan, RADAR: An In-Building RF-Based User Location and Tracking System, In *Proc. of the IEEE INFOCOM 2000,* Tel-Aviv, Israel, March 2000.
2. S. Basagni, I. Chlamtac, V. Syrotiuk, and B. Woodward, A Distance Routing Effect Algorithm for Mobility (DREAM), In *Proc. of MOBICOM 1998,* Dallas, TX, USA, October 1998.
3. S. Brands and D. Chaum, Distance-bounding protocols, in Workshop on the Theory and Application of Cryptographic Techniques on Advances in Cryptology, Springer-Verlag New York, Inc., 1994, pp. 344-359.

[5] This document was prepared through the collaborative participation in the Communication and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. The views and conclusions contained in this documents are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

4. N. Bulusu, J. Heidemann and D. Estrin, GPS-less Low Cost Outdoor Localization for Very Small Devices, In *IEEE Personal Communications Magazine*, 7(5):28–34, October 2000.
5. S. Čapkun and J.-P. Hubaux, Secure Positioning in Wireless Networks, IEEE Journal on Selected Areas in Communications, vol. 24, no. 2, February 2006.
6. S. Čapkun, L. Buttyan, J. Hubaux, SECTOR: Secure Tracking of Node Encounters in Multi-hop Wireless Networks, in *Proc. of SASN 2003*, Fairfax, Virginia, October 2003.
7. S. Čapkun, M. Cagalj, and M. Srivastava, Secure Localization with Hidden and Mobile Base Stations, in Proceedings of the IEEE Conference on Computer Communications (InfoCom), 2006.
8. D. Coppersmith and M. Jakobsson, Almost optimal hash sequence traversal, In *Proc. of the FC 2002*, Lecture Notes in Computer Science, IFCA, Springer-Verlag, Berlin Germany, 2002.
9. N. Cressie, *Statistics for Spatial Data*, John Wiley & Sons, 1993.
10. L. Doherty, L. Ghaoui and K. Pister, Convex Position Estimation in Wireless Sensor Networks, In *Proc. of the IEEE INFOCOM 2001*, Anchorage, April 2001.
11. J. Douceur, The Sybil Attack, In *Proc of IPTPS 2002*, Cambridge, MA, USA, March 2002.
12. T. He, C. Huang, B. Blum, J. Stankovic and T. Abdelzaher, Range-Free Localization Schemes in Large Scale Sensor Network, In *Proc. of MOBICOM 2003*, San Diego, CA, USA, September 2003.
13. Y. Hu, A. Perrig, and D. Johnson, Packet Leashes: A Defense Against Wormhole Attacks in Wireless Ad Hoc Networks, In *Proc. of INFOCOM 2003*, San Francisco, CA, USA, April 2003.
14. M. G. Kuhn, An Asymmetric Security Mechanism for Navigation, in Proceedings of the Information Hiding Workshop, 2004.
15. L. Lamport, Password Authentication with Insecure Communication, In Communications of the ACM, 24(11):770-772, November 1981.
16. L. Lazos and R. Poovendran, Energy-Aware Secure Multicast Communication in Ad-hoc Networks Using Geographic Location Information, In *Proc. of IEEE ICASSP 2003*, Hong Kong, China, April 2003.
17. L. Lazos and R. Poovendran, SeRLoc: Robust Localization for Wireless Sensor Networks, ACM Transactions on Sensor Networks (TOSN), August 2005, vol. 1, pp. 73–100.
18. L. Lazos and R. Poovendran, HiRLoc: High Resolution Localization for Wireless Sensor Networks, IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Network Security, February 2006, Vol. 24 (2), pp. 233–246.
19. L. Lazos and R. Poovendran, SeRLoc: Secure Range-Independent Localization for Wireless Sensor Networks, ACM Workshop on Wireless Security, October 2004, (WiSe '04), pp. 21–30.
20. L. Lazos, S. Čapkun and R. Poovendran, ROPE: Robust Position Estimation in Wireless Sensor Networks, 4th International Symposium on Sensor Networks, April 2005, (IPSN '05), pp. 324–331.
21. Z. Li, W. Trappe, Y. Zhang, and B. Nath, Robust Statistical Methods for Securing Wireless Localization in Sensor Networks, in Proceedings of the International Conference on Information Processing in Sensor Networks (IPSN), 2005.
22. D. Liu and P. Ning, Location-based pairwise key establishments for static sensor networks, In *Proc. of SASN 2003*, Fairfax, VA, October 2003.
23. D. Liu, P. Ning, and W. Du, Attack-Resistant Location Estimation in Sensor Networks, in Proceedings of the International Conference on Information Processing in Sensor Networks (IPSN), 2005.

24. MICA Wireless Measurement System, available at: http://www.xbow.com /Products/Product_pdf_files /Wireless_pdf/MICA.pdf.
25. R. Nagpal, H. Shrobe and J. Bachrach, Organizing a Global Coordinate System from Local Information on an Ad Hoc Sensor Network, In *Proc. of IPSN 2003,* Palo Alto, USA, April, 2003.
26. J. Newsome, E. Shi, D. Song and A. Perrig, The Sybil Attack in Sensor Networks: Analysis and Defenses, In *Proc, of IPSN 2004*, Berkeley, CA, April 2004.
27. D. Nicolescu and B. Nath, Ad-Hoc Positioning Systems (APS), In *Proc. of IEEE GLOBECOM 2001,* San Antonio, TX, USA, November 2001.
28. P. Papadimitratos and Z. J. Haas, Secure Routing for Mobile Ad Hoc Networks, in *Proc. of CNDS 2002*, January 2002.
29. A. Perrig, R. Szewczyk, V. Wen, D. Culler and J. Tygar, SPINS: Security Protocols for Sensor Networks, In *Proc of MOBICOM 2001,* Rome, Italy, July 2001.
30. R. Poovendran and L. Lazos, A Graph Theoretic Framework for Preventing the Wormhole Attack in Wireless Ad Hoc Networks, to appear in ACM/Springer Journal on Wireless Networks (WINET).
31. N. Priyantha, A. Chakraborthy and H. Balakrishnan, The Cricket Location-Support System, In *Proc. of MOBICOM 2000,* Boston, MA, USA, August 2000.
32. R. L. Rivest, The RC5 encryption algorithm, In *Proc. of the first Workshop on Fast Software Encryption*, pp. 86-96, 1995.
33. N. Sastry, U. Shankar and D. Wagner, Secure Verification of Location Claims, In *Proc. of WISE 2003,* San Diego, CA, USA, September 2003.
34. A. Savvides, C. Han and M. Srivastava, Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors, In *Proc. of MOBICOM 2001,* Rome, July 2001.
35. Y. Shang, W. Ruml, Y. Zhang and M. Fromherz, Localization from Mere Connectivity, In *Proc. of MOBIHOC 2003,* Annapolis, MD, USA, June 2003.
36. B. Hofmann-Wellenhof, H. Lichtenegger and J. Collins, Global Positioning System: Theory and Practice, Fourth Edition, Springer-Verlag, 1997.
37. D. Stinson, *Cryptograhpy: Theory and Practice*, 2nd edition, CRC Press, 2002.
38. R. Pickholtz, D. Schilling, and L. Milstein. Theory of Spread Spectrum Communications - A Tutorial, In the *IEEE Transactions on Communications*, 30(5):855–884, May 1982.
39. S. B. Wicker and M.D. Bartz, Type-II Hybrid-ARQ Protocols Using Punctured MDS Codes, In *Proc. of IEEE Transactions on Communications*, April 1994.